R E P O R T R E S U M E S

ED 016 294 24 EA 001 075

A MEASUREMENT OF COLLEGE INSTRUCTOR BEHAVIOR. SUMMARY.

BY- FAHEY, GEORGE L.
PITTSBURGH UNIV., PA.

REPORT NUMBER BR-6-8047 PUB DATE 15 MAR 67
CONTRACT OEC-6-10-355
EDRS PRICE MF-$0.25 HC-$0.36 7P.

DESCRIPTORS- *COLLEGE TEACHERS, *MEASUREMENT TECHNIQUES, *TEACHER BEHAVIOR, STUDENT OPINION, BEHAVIOR RATING SCALES, *EFFECTIVE TEACHING, LIBERAL ARTS, COLLEGE STUDENTS, HUMANITIES, NATURAL SCIENCES, SOCIAL SCIENCES, *TEACHER CHARACTERISTICS, STATISTICAL ANALYSIS, PITTSBURGH,

THE PROBLEM OF ACHIEVING DELINEATION OF COLLEGE TEACHING BEHAVIOR IS INVESTIGATED. THE STUDY HAS FOUR OBJECTIVES--(1) TESTING THE EFFECTIVENESS OF A SEMANTIC DIFFERENTIAL SCALE AS A MEANS OF DESCRIBING INSTRUCTOR BEHAVIORS AS OBSERVED BY STUDENTS, (2) TESTING THE EFFECTIVENESS OF CERTAIN SINGLE ADJECTIVES PURPORTED TO BE DESCRIPTIVE OF INSTRUCTOR BEHAVIORS, (3) UTILIZING OBTAINED RATINGS OF INSTRUCTORS TO DIFFERENTIATE BETWEEN STUDENTS' VIEWS OF INSTRUCTORS IN DIFFERENT MAJOR AREAS OF ACADEMICS, AND (4) RELATING OBTAINED DESCRIPTIONS OF INSTRUCTORS TO JUDGED EFFECTIVENESS. THIRTEEN SINGLE ADJECTIVES, 12 BIPOLAR PAIRS, AND ONE SINGLE ITEM GLOBAL JUDGMENT OF TEACHING EFFECTIVENESS COMPRISED A RATING SCALE WHICH WAS ADMINISTERED TO 4,916 UNDERGRADUATE STUDENTS IN THE CLASSES OF 116 LIBERAL ARTS INSTRUCTORS IN HUMANITIES, NATURAL SCIENCES, AND SOCIAL SCIENCES. FOUR MAJOR CONCLUSIONS WERE REACHED--(1) COLLEGE STUDENTS CONTINUE TO BE RELIABLE IN THEIR DESCRIPTIONS OF INSTRUCTORS, (2) ONLY THE ADJECTIVE "STIMULATING" RELATED HIGHLY TO THE GLOBAL JUDGMENT OF EFFECTIVENESS SO THAT FURTHER RESEARCH WITH ADJECTIVE LABELS FOR BEHAVIOR MAY BE FUTILE, (3) THE SEMANTIC DIFFERENTIAL SCALE CONTRIBUTED NOTHING TO THE DESCRIPTION OF INSTRUCTOR BEHAVIOR, AND (4) WIDE VARIANCE WAS EVIDENCED BETWEEN INDIVIDUAL INSTRUCTORS ON ALL THE SEPARATE ITEMS OF THE SCALE INCLUDING THE GLOBAL JUDGMENT OF EFFECTIVENESS. (HW)

ED016294

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

Summary

A MEASUREMENT OF COLLEGE INSTRUCTOR BEHAVIOR

George L. Fahey

University of Pittsburgh

Project Number 6-8047 PA 24

15 March, 1966 - 15 March, 1967

EA 001 075

BACKGROUND

Adequate descriptions of college teaching behaviors would facilitate selection, placement, evaluation, and development of instructors. Numerous attempts at such description have not yielded convincing results. The means of achieving delineation in the portrayal of college teaching behavior has not as yet been developed. This study has investigated other approaches to the problem.

OBJECTIVES

1. To test the effectiveness of a Semantic Differential Scale as a means of describing instructor behaviors as observed by students.
2. To test the effectiveness of certain single adjectives purported to be descriptive of instructor behaviors.
3. To utilize obtained ratings of instructors to differentiate between students' views of instructors in different major areas of academics: natural sciences, social sciences, and humanities.
4. To relate obtained descriptions of instructors to judged effectiveness.

PROCEDURE

Single adjectives from a list proposed by Fahey (1963) and Ryans (1960) and bi-polar pairs of adjectives from the Semantic Differential Scale by Osgood, Suci, and Tannenbaum (1951) were presented in a pilot study to 116 undergraduate and graduate students who rated three college and University instructors anonymously. Obtained ratings were factor analyzed and three factors identified.

The pilot study showed that the eight approaches to college teaching proposed by Fahey (1963) were excessively subject to a common variance.

Ratings on one approach being virtually identical with those on each other approach, all were consolidated into one "approach to teaching". This deviation from the original proposal seemed fully warranted. Thirteen single adjectives and 12 bi-polar pairs having highest loadings on their respective factors were retained in a rating instrument. A single item requesting a global judgment of teaching effectiveness was added.

This rating scale was administered to 4,916 undergraduate students in the classes of 116 liberal arts instructors categorized as in the academic divisions of humanities, natural sciences, or social sciences. The measure was applied in a second semester to 2,967 undergraduates in the classes of 59 instructors of the original sample. Instructor participation at either testing was voluntary.

Obtained data were used to determine reliability of the rating instrument. A discriminate analysis was used to test any differentiation between the academic divisions. Each adjective and bipolar pair was correlated with each other including the global judgment item used as a criterion of effectiveness.

RESULTS

1. The rating scale was found to have high reliability. Out of a total of 78 coefficients, the lowest was .85, next lowest was .88, and the remaining 76 exceeded .90. The reliability of the criterion item of effectiveness also exceeded .90.
2. Discriminant analysis showed that the three academic division samples should be considered as one sample. That is, descriptions of instructors afforded by the instrument did not differ in relation to whether they taught humanities, natural sciences, or social sciences.

3. Only one of the single adjectives, "stimulating", related well with the global judgment of effectiveness. None of the other 12 was found to have a relationship as large as .50.
4. Three of the 12 bipolar pairs of adjectives were correlated with the criterion to the extent of .50 or more: "good-bad", .65; "valuable-worthless",.61; "clear-hazy", .54. Since two of these pairs essentially ask the same question as the criterion item, only "clear-hazy" appeared to afford information about instruction.
5. Factor analysis of the obtained data extracted six factors. Three of these factors contributed appreciably to an explanation of variance between high and low ratings on the effectiveness criterion. The factor named "Professional Impact" explained 31 percent of the variance. It included the bipolar pairs clear-hazy, good-bad, valuable-worthless, near-far, and the single adjectives vague, stimulating, planful. The factor named "Abrasive Impact" explained 23 percent of the variance. It included the single adjectives: egotistical, cynical, rude, pretentious, rebellious and no bipolar pairs. The third factor named "Sensory Impact" explained 18 percent of the variance. It included the bipolar pairs, bass-treble, wet-dry, thick-thin, loud-soft, and no single adjectives. The other three factors extracted each explained less than 10 percent of the variance.

CONCLUSIONS

1. As in previous research findings, college students continued to be reliable in their descriptions of instructors. The high reliability coefficients for all scale items showed remarkable

consistency or like-mindedness among students as they reported on individual teachers. Inconclusiveness in the findings must be explained on grounds other than the capacity of students to make consistent reports.

2. The list of 13 single adjectives included only one which related highly to the global judgment of effectiveness. This adjective, "stimulating", has been evidenced in other researches. Some other adjectives did not singly but did in patterns appear to have validity, as with the "abrasive impact" factor. It may be futile to attempt further research with adjectival labels for behavior.

3. The Semantic Differential Scale contributed nothing to the description of instructor behavior in this study. Two bipolar pairs stood up but essentially asked the criterion question. The other pair which correlated favorably is almost as obviously a variant of the criterion item. A general inditement of the Semantic Differential tool for research is not warranted from these findings. The instrument was designed to measure meaning. It has been used elsewhere to describe behavior and, from the present findings, is not appropriately used for such a purpose. The inference that a rating is a projection of the meaning of the observed to the observer is either untenable or the instrument cannot detect such meaning.

4. While this study did not demonstrate differences between patterns of instructor behavior in major academic divisions, it did show evidence of wide variance between individual instructors on all the separate items of the scale including the global judgment of effectiveness. Analysis of this variance is being continued beyond the present project.

BIBLIOGRAPHY

There are 136 references listed in the dissertation of which this project was a part.

PUBLICATIONS

Apt, Madeline H. "A Measurement of College Instructor Behavior". Unpublished Ph.D. dissertation, University of Pittsburgh, 1966.

Apt, Madeline H. "A measurement of college instructor behavior. J. of Educational Measurement. Submitted. (This is submitted as the Final Report of this Project.)